# AUDITORY TESTS.[1]

BENJAMIN RICHARD ANDREWS.

TABLE OF CONTENTS.


| | | |
|---|---|---|
| Chapter IV. | Tests of Musical Capacity | 302 |
| | Introduction | 302 |
| | I. Discrimination of Tones | 303 |
| | a. With tone-tester | 304 |
| | b. With forks | 305 |
| | c. Other methods | 306 |
| | II. Discrimination of Chords | 307 |
| | a. Recognition of harmonic intervals | 307 |
| | b. Unitariness of harmonic intervals | 308 |
| | III. Rhythm | 309 |
| | a. Involuntary rhythmisation | 311 |
| | b. Rhythmisation with suggestion | 312 |
| | c. Musical phrases and sentences | 314 |
| | IV. Affective Reaction to Music | 314 |
| | a. Harmonic intervals | 314 |
| | b. Chords | 315 |
| | c. Musical compositions | 316 |
| | V. Additional Tests and Questions | 316 |
| | Summary of Chapter IV. | 318 |
| Chapter V. | Diagnostic Tests of Hearing | 318 |
| | I. Speech Tests | 319 |
| | II. Integrity of Tonal Scale | 320 |
| | III. Determination of Deafness | 321 |
| | IV. Binaural Pitch Difference | 322 |
| | V. Special Tests: | 322 |
| | Weber's, Rinne's, Gellé's, Schwabach's | 323 |
| General Summary | | 325 |


## CHAPTER IV.

## TESTS OF MUSICAL CAPACITY.

The ability to appreciate music cannot be measured simply and directly, since musical perception is a conscious experience of wide range and very great complexity. It remains to approach the problem indirectly, through the examination of the more fundamental factors which contribute to the apprehension and enjoyment of musical form. Music involves a sequence of single notes (melody), or more usually a sequence of tonal complexes (harmony), possessing in either case rhythmical and metrical characteristics, and evoking an affective or emotional

[1] Continued from Vol. XV, January, 1904, p. 14.

response on the part of the hearer. In musical perception, we may accordingly distinguish four factors, an examination of which will afford a view, even if but a partial view, of the musical capacity of an individual:

1. The discrimination of single notes heard in succession.
2. The discrimination of groups of simultaneous notes (chords) heard in succession.
3. The perception of rhythm.
4. The emotional or affective response to music.

Our musical tests fall then, naturally, into four groups: pitch discrimination, chord discrimination, rhythm, and affective response. To these is added a fifth group of miscellaneous tests and methods. The paper is intended to present a systematic view of the whole field; not a programme necessarily to be followed in its entirety in individual cases of testing.

I. *Discrimination of Tones.* The problem, here, is to measure the individual's capacity for discriminating successive tones of different pitch. The test will consist in submitting to the observer a single tone of a certain pitch, followed, after a brief interval, by a second tone of different pitch, and in requiring judgment as to the pitch relations of the two tones. This problem has been approached experimentally by Gilbert[1] and by Seashore,[2] the former employing an adjustable pitch pipe, the latter, a series of tuning forks. Their methods will be set forth presently.

Several preliminary matters must first be disposed of. (1) In any judgment of pitch, a difference in intensity may be mistaken for a difference in pitch. This is to be avoided by keeping the intensity of the successive sounds as nearly as possible uniform.[3] (2) Stern has shown that certainty of judgment is affected by the length of the time interval between the two tones.[4] He names six seconds as optimal, and though a shorter period, 2-3 seconds, may be used, it should at any rate be kept uniform. (3) More important is the form of judgment required. The subject may be asked to state each time whether the second tone is "like" the first, or "different;" or whether the second tone is of the same pitch as the first, or is higher or lower than it. Gilbert used the former and Seashore the latter form of judgment. It is recognized that at the limen it is much easier to judge whether two tones are alike or different in pitch,

[1] J. A. Gilbert: *Experiments on the Musical Sensitiveness of School Children, Studies from Yale Psych. Lab.*, I, 80.

[2] C. E. Seashore: *Hearing Ability and Discriminative Sensibili for Pitch, Univ. of Iowa Studies in Psych.*, ii, 55.

[3] C. Stumpf, *Tonpsychologie*, 1883, I, 315, mentions a case in which judgments of pitch-difference were always made in terms of intensity. The louder tone was always judged to be higher.

[4] L. W. Stern: *Zeitsch. f. Psych. u. Phys.*, xxi, 1899, 377-379.

than to state particularly whether the one tone is lower or higher in pitch than the other. In the one case, the judgment is of simple difference; in the other, of qualitative difference of direction.[1]

The judgment of qualitative difference is to be preferred, since (a) it tests directly the sort of sensible discrimination required in singing or playing, or in listening to music—in such cases one is interested in the rising and falling of notes, and not in the existence of simple difference; and (b) it is more difficult, and hence gives a more rigid test than discrimination of simple difference.[2]

We may now consider various tests of pitch discrimination.

A. *Determination with Tone-tester.* The instrument employed is an adjustable pitch-pipe, or "tone-tester," which is fully described by Gilbert,[3] who used it in his tests. It has a range over the octave from $f^1$ to $f,^2$ and is scaled to test pitch discrimination at $a^1$ of 435 vibrations (in some pipes, 450 vibrations). The tester is mounted at the base of a fan-shaped board and an index arm continues its adjusting lever to a scale on which are subdivisions for thirty-seconds of a tone, over a half-tone above and below $a^1$. Uniform intensity of sound can be approximately secured by using the residual air in the lungs after normal expiration. The instrument is especially suitable for mass-tests, as in a schoolroom, and in such use judgments may be written.

The method to be employed with the tone-tester is that of minimal changes, 'with knowledge.' The following series of pairs of tones is submitted: a′ and a′, a′ and a′ + 1/32, a′ and a′ + 2/32, . . . . . and this is continued to some pitch of a′ + n/32 such that it lies well above the point at which the second tone is always judged to be of "higher" pitch than the first.

---

[1] W. Preyer: *Grenzen d. Tonwahrnehmung*, 1876, pp. 2, 36; Stumpf: *op. cit.*, I, 313.

[2] The writer computed Seashore's and Gilbert's results to see if differences could be found which might be ascribed to the use of the two different forms of judgment. Seashore's work on adults (p. 56), excluding six who could not recognize a difference of a semi-tone, gives an average discrimination of .142 of a tone for 18 women, and of .173 of a tone for 25 men (all college students). Gilbert, working with children from 12 to 19 years old, found an average discrimination of .112 of a tone. This seems to indicate the influence of the two sorts of judgment used, as the limen falls lower in Gilbert's work, which was based on the easier judgment of simple difference. The age difference in the two cases does not affect this conclusion, since the limen decreases with age (Gilbert), and here the younger subjects have the lower limen. The different instruments employed may, in some measure, account for the discrepancy in the limens obtained in the two investigations; but doubtless the different form of judgment is the real explanation.

[3] Gilbert: *op. cit.*, 81.

The "ready" signal is given two seconds before each pair to ensure maximal attention. The series 'is then reversed in order, beginning at this upper value a′ + n/32, and a descending series given: a′ and a′ + n/32, a′ and a′ + n—1/32, etc.[1]

The subject should be informed which way the changes are to progress. The pitch in the ascending series which is first judged "higher," and the pitch in the descending series which is last judged "higher" before passing to the judgment of "equal," are averaged for the first result. Five or ten tests should be made and the results averaged.[2] The mean variation should be computed as a check on errors.

B. *Method with tuning forks.* This method, which employs a standard fork of a′ (435 vibrations) and a series of forks differing from it by from 1 to 30 vibrations, has been used by Seashore.[3] The series of forks includes, in addition to the primary a′ fork, nine secondary forks whose vibration rates are respectively 1, 2, 3, 5, 8, 12, 17, 23 and 30 vibrations higher than 435. At this point on the scale, a difference of 54 vibrations is equivalent to a difference of a whole tone. Accordingly, the tone of the primary fork followed by that of the fork one vibration higher gives a pitch difference of 1/54 of a tone; and with the other forks it gives pitch differences of 2/54, 3/54, 5/54, etc., of a tone. The forks should be set in vibration by striking them against a rubber cushion, or the palm of the hand, and with care to secure as nearly as possible uniform intensity. The forks should always be held in the same position before the ear; the best position is just before the meatus, with the plane of the two forks parallel to the side of the head. The standard tone is sounded first, and judgment required as to whether the second tone is the same as, or higher than, the first. The method is like that with the tone-tester—the average is taken of an ascending and descending series of minimal

[1] Gilbert uses, as his second series, a descending series below a′: thus, a′ and a′, a′ and a′—1/32, a′ and a′—2/32, . . . . . He then averages the results of the ascending series above a′ and the descending series below a′. This involves determination of the limen at two distinct points on the scale, and the averaging of these unrelated values. The method recommended above obviates this by finding values in both ascending and descending series for the same liminal band on the scale. In tests of this degree of fineness, the two methods would scarcely give different results. The distinction between them, however, should be kept clear, and it is advisable to keep to the more accurate procedure.

[2] Gilbert found an average discrimination of a little more than 5/32 of a tone. He correlates discrimination with age and finds a fairly regular increase from .384 of a tone at 6 years to .075 of a tone at 19 years.

[3] Seashore: *op. cit.*, 59, and *Educ. Rev.* xxii, 1901, 75.

changes with knowledge,[1] and then the determination is repeated five or ten times.[2]

C. *Other tests of Pitch Discrimination.* The Tests Committee[3] of the American Psychological Association makes some suggestions regarding tests of perception of pitch.[4] Cattell and Jastrow recommend the adjustment of "one monochord or pipe to another, the tones not to be sounded simultaneously;" and Jastrow suggests the use of two Gilbert tone-testers for this purpose. Baldwin and Sanford recommend the selection of "a match from a set of forks, giving a fixed number of vibrations per second more or less than a standard, *e. g.*, a standard fork of 500 vibrations per second, and other forks 497, 497.5, 498, etc., 500, 500.5, 501, 501.5, etc." The Committee suggests no plan of procedure. The intention is apparently to put the manipulation of instruments in the hands of the persons tested. This procedure would open the way to variable errors from differences in the intensity of the tones used, the position of the forks before the ear, etc. In any case, a single determination would furnish no measure of pitch discrimination. Two tones are subjectively equal if the second falls anywhere within the limits of indiscriminability of difference above and below the first tone. A single adjustment to subjective equality might accordingly by chance hit upon objective equality, or might exist with an objective difference in pitch of any amount up to the limen. Tests which involve

[1] Seashore's method is somewhat different, and it may be followed instead of the one suggested. Starting with the widest interval (made by the forks $a'$ and $a' + 30$ vib. sounded successively) he gives one trial at each step until the region of uncertainty is reached; then ten trials at each step. He takes as the limen the lowest interval for which there are eight correct judgments in the ten tests made. The method recommended has an advantage over this of Seashore, in that the averaging of results from ascending and descending series provides against the error of habituated judgment.

[2] On the basis of tests on 110 adults and 380 children, Seashore suggests tentatively the following scale for the interpretation of results: (1) a child whose limen is 2/54 of a tone or less may become a musician; (2) 3/54-8/54 should have a plain musical education, and singing in school should be obligatory; (3) 9/54-17/54, singing in school optional, and a plain musical education should be given if special inclination for music is shown; (4) 18/54 or above, should have nothing to do with music (*Educ. Rev.*, xxii, 76). This scale is not only arbitrary, but its application would result in injustice. It is not fair to draw final conclusions as to musical capacity from tests on a single factor, as pitch discrimination. Moreover, pitch discrimination depends in good part on practice: a person who could just discriminate a difference of 20/54 of a tone might, with practice, reduce this to 10/54. For such a case, see Seashore's own results, *Univ. of Iowa Studies*, II, 64.

[3] *Cf.* preceding portion of article, this *Journal*, Jan., 1904, 23.

[4] *Psych. Rev.*, iv, 1897, 134.

"matching" or "adjusting" to equality, can, therefore, only be carried out by the method of average error; *i. e.*, they require a series of adjustments, 10 or more, and are measured by the average of the absolute deviations (regardless of + or —) from objective equality. Still the errors due to the subject's handling the instruments persist. Accordingly the two methods outlined are, perhaps, to be preferred to those recommended by the Committee.

It seems especially desirable to suggest some method which can be used for rough and rapid work, in case the instruments recommended are not at hand. It is quite possible with the piano or organ to make tests which will disclose the presence of extremely poor pitch discrimination, though, of course, no measurement of the degree of discrimination is possible. The person to be tested should be placed so that he cannot see the key-board and care should be taken to strike the keys with equal force, and to allow an equal interval between the members of each pair of notes. Such a series as the following should be submitted, the "ready" signal being given before each pair of tones, and the tones presented in succession: c′—c′; c′—e′♭; g′—g′♯; g′—g′; g′—f′♯; d′—d′♯; d′—d′; e′—f′; c′—c′♯; g′—g′♯; g′—g′. The person is to judge whether the second tone is the same as, or higher or lower than, the first. In this way, it can be determined in a few moments whether a person always recognizes correctly pitch differences of a semi-tone or more. The tests may be repeated in the octave above and below the middle octave of the piano, to determine whether differences in pitch are equally well perceived with high and low tones. Such a test admits of simultaneous application to a large number of individuals, as in a schoolroom. Rough tests may also be made for intervals less than a semi-tone by using an unmounted adjustable pitch pipe, like that adapted for use in the tone-tester. By setting its adjusting arm for small differences, especially where the interval on the adjusting scale is large, it is possible to sound successive tones varying by 1/2 or 1/3 a semi-tone, *i. e.*, 1/4 or 1/6 a tone. In this way, with the piano and the adjustable pitch pipe, a rough examination of capacity in pitch discrimination can be made.

II. *Discrimination of Chords.* Another factor involved in musical capacity is the perception of simultaneous tonal combinations. Such combinations, as ordinarily employed in music, are composed of three or more simultaneous notes and are called chords; combinations of two simultaneous notes form harmonic intervals. The latter lie at the basis of the former, and, for convenience' sake we may test them rather than the more complex combinations. Tests of two kinds are suggested:

(A) The perception of identity or difference in harmonic intervals given successively.

(B) Judgment as to the simplicity or complexity of single harmonic intervals given without knowledge of their objective tonal composition.

A. *The recognition of successive harmonic intervals.* This test measures the recognition of likeness and difference in successively perceived harmonic intervals. The person is asked

to pass judgment of "like" or "different" on pairs of intervals successively given on a piano or organ. A series like the following may be used: 5th—5th; 5th—octave; major 3rd—minor 3rd; major 3rd—major 6th; minor 3rd—minor 6th; major 2nd—minor 2nd; 5th—major 3rd; 5th—minor 6th; 5th—5th; 5th—minor 3rd; major 7th—minor 7th; octave—octave; octave—major 7th; etc. The two intervals compared must always have a common note. The results may be expressed in percentages, and the individuals' results referred to the average result.

B. *Unitariness or Plurality in Harmonic Intervals.* This test involves judgment as to unitariness or plurality, or as to the degree of unitariness, in harmonic intervals submitted without knowledge of their composition. Stumpf includes this as one of his four tests to determine unmusical ears.[1] This test takes into account the fact that the constituents of an interval are heard out with a facility which varies according to the interval involved, and which further varies directly with the musical ability of the person tested. The latter fact has given rise to two methods which we will state, though the first alone is practicable for inexpert use, as in tests. The average untrained observer, under certain conditions of experimentation, mistakes the octave for a single tone in about 75% of his judgments; the 5th, 40-60%; the 4th, 28-36%; 3rd's and 6th's, 20-30%; 2nd's and 7th's, about 15%[2]. This may be expressed by saying that the octave is the closest fusion, and that the other intervals are less closely fused, in the order given.

(a) The first test of perception of degree of fusion is used in psychological determinations for untrained, unmusical subjects, and is the only one which can ordinarily be used in tests. It consists of submitting various harmonic intervals, one by one, to the person tested, and requiring a judgment whether one note or two notes are heard. In most instances it will scarcely be possible to go through a sufficiently long series to warrant comparative percentage results for all the intervals. It is suggested that 30 or more intervals be submitted for judgment, including octaves ($c^1$—$c^2$), 5th's ($c^1$—$g^1$), 3rd's ($c^1$—$e^1$ and $c^1$—$e^{1\flat}$), and 2nd's ($c^1$—$d^1$ and $c^1$—$d^{1\flat}$). These intervals should be arranged in an irregular order and single notes interpolated among them as a check on judgments given; thus: $c^1$, $c^1$—$g^1$, $c^1$—$c^2$, $g^1$, $c^1$—$d^1$, $c^1$—$e^1$, $e^1$, $c^1$—$c^2$. . . . At each step the subject

[1] C. Stumpf: *Tonpsych.*, ii, 157. The other tests are: singing a note which has been played on the piano; judgment as to which of two successive tones is the higher; and judgment as to the more pleasant of two successively given fusions.

[2] C. Stumpf: *ibid.*, 135, 140, 142 ff.; the conditions are stated in E. B. Titchener's *Experimental Psychology*, Vol. I, Pt. I, 330 ff.

will state whether he hears a single note or two notes. It will thus be possible to determine how the person perceives the intervals of high, medium, and low degree of fusion. One may expect mistakes to occur frequently, perhaps more than half the time, in judgments of the octave; less frequently with the 5th's and 3rd's; and only occasionally with the 2nd's. If results vary considerably from the values given above, more extended tests may be made. Individual results may be roughly referred to the percentages given, or the average results found.

(b) In case the person tested is musical, another method, the direct method of paired comparison, may be used. In this method, the various intervals are compared, each with every other, and comparative judgment passed directly as to which is the better fusion. In such a judgment, the pleasantness or unpleasantness of the impressions, and considerations of musical significance, and of nearness of the constituent tones in the scale, must be abstracted from. The certainty with which one interval is judged a better fusion than another should be noted, as this is an important indication of degree of unitariness. The results as before may be compared with the percentage scale given above, or each result compared with the average. This method is, however, little suited for unskillful use.

In both methods the piano or organ will usually be found most convenient. The first method can be readily adapted to school or other mass tests.

By these two tests, first the recognition of likeness or difference in successive harmonic intervals, and secondly, the perception of unitariness of harmonic intervals, reliable indications of musical capacity may be obtained, so far as this capacity depends upon correct perception and discrimination of tonal complexes.

III. *Tests of the Perception of Musical Rhythm.* As noted in the introduction, music is not simply a composition of simultaneous and successive tones, but it possesses rhythmical and metrical characteristics. A musical composition, as a progress in time, is divided into equal temporal units called measures, and into larger divisions, composed of groups of measures, called "phrases," "sentences," "periods," etc. The measure is indicated in printed music as the group of notes falling between two perpendicular lines or bars on the staff. The conception of the measure involves two factors: the durations of the pitches which "fill out" the measure; and the stresses or accents which the notes receive. In the first place, the temporal sum of the notes of successive measures is the same. In the second place, the pattern or scheme of accents or stresses remains the same in successive measures—a heavy stress on the initial note

of the measure and one or more weak stresses on the following notes in the measure. A measure extends from one heavy stress to the next. There are two fundamental rhythmical patterns: that of two accents to the measure, the first strong, the second weak, hence known as 2-part or duple rhythm, march rhythm and common time; and that of three accents to the measure, the first strong, the last two weak, hence known as triple rhythm or waltz time. There are possible many accentual patterns, but all are reducible to these two fundamental forms. The same pattern usually continues throughout the same composition. A musical composition is thus a succession of groups of notes, or measures, each conforming to the accentual pattern of the piece. The equal duration and the accentual similarity of successive measures, are the primary reason for music appearing as a temporal movement with related parts, that is, as a rhythm.

In addition to the rhythm dependent upon similarity between measures, there is a larger rhythm in music based upon the stressing of the larger structural units, the "section," "phrase," "sentence," and "period." These terms are used variously, but they indicate groups of measures, composed of two, four, eight, or more measures, as the case may be. The smaller groups are component parts of the larger groups: two of the 2-measure groups follow each other and make a 4-measure group; two of the 4-measure groups make an 8-measure group, etc. The initial part of a large group is stressed as is the initial note of a measure, and major and minor accents, respectively, set off the larger group of measures from one another, and indicate their subdivisions or the smaller measure-groups within them. There are thus accentual patterns having reference to the phrase, sentence, etc., just as the 2-part or 3-part pattern has reference to the measure. So there is a larger rhythm in terms of these groups of measures.

The tests[1] suggested refer first to the rhythm directly depend-

[1] There is some experimental evidence of individual variations in the sense of rhythm. Sears, *Ped. Sem.*, viii, 1901, 19, secured information by a questionary of several hundred children who showed difficulty in keeping time in the movements of marching, dancing, and calisthenics. He refers to statements of teachers, army officers and dancing masters, that there are wide variations in the ability of individuals to appreciate the rhythm of music and respond to it with proper movements. Some cases reported were deficient in rhythmical sense and remained so despite training; others, who had been deficient, had improved under training. In his experimental work, however, Sears found only 9 out of 1,297 children who were entirely unable to repeat by tapping objective rhythmical series set them. Bolton (*Amer. Jour. of Psych.*, vi, 1894, 185 and 204) found a larger percentage, 2 out of 50 persons examined, who were quite deficient in rhythm. His method was to submit a series of uniform metronome clicks to his subjects

ent upon similarly accented measures; second, to the larger rhythm of the phrase, sentence, etc. Perception of rhythm of the first sort can best be examined with non-musical sounds, and there are suggested: (A) Tests for *involuntarily* hearing a succession of non-musical sounds as a rhythm series; (B) Tests for hearing such a succession as a rhythm, *voluntarily or with suggestion.* To this section are added certain tests for hearing, as a rhythm, a succession of musical notes, in which variations in duration, quality and intensity occur; and, also, a test for recognizing the fundamental two-part and three-part accentual patterns of musical measures. (C) Tests are suggested for the larger rhythm referred to, the perception of musical phrases and sentences.

A. *Tests of Involuntary Rhythmisation.* The test used by Bolton, referred to above, is to be employed here. It consists in submitting to the subject a series of successive sounds which are entirely identical and follow each other at a regular brief interval; the subject is to report how he hears them, *i. e.*, whether as entirely similar, disparate sounds, or whether he subjectively arranges them into accented rhythmical groups.

The details of such a test are as follows: A metronome[1] is used which gives clicks of the same intensity and quality, and is adjustable to various rates, from 40 to 208 per minute. The bottom should be removed from the base of the metronome and it should be set on a pad of heavy cloth or felt, to deaden the sounds. One of the more rapid rates, 152 or 200 beats per minute, should be used, as this is more favorable for subjective rhythmisation, and it should be continued for 45 seconds. The following general directions should be given the person tested: he is to put himself in a receptive and passive attitude toward the sounds; he is not to fix his attention closely on each sound as such, but simply listen to the successive beats and tell how he hears them, whether he notices anything peculiar about them. Any reference to rhythm, or to the fact that this is a test of rhythm must be avoided. The first series should be repeated three times; and then other repeated series should be given, in which the 200, 152 and 92 rates at least should be included.

With some persons, rhythmical groupings of the uniform sounds will take place almost unconsciously, and manifest themselves in some way; perhaps, in an involuntary counting, "one—two, one—two," or in other accented verbal accompaniment; perhaps, in movements in time of hand, head or foot; or by some accompanying imagery, auditory or visual, as the ticking of a clock, a mental picture of a person walking, the

---

and only two failed to group them in some rhythmical order. It is important to note that these two persons "possessed no appreciation of music at all, they could not carry a tune." Yet they must have had some notion of music for "they were able to recognize some of the common airs when they were sung or whistled."

[1]Titchener: *op. cit.*, Vol. I, Pt. 1, 174, and Pt. 2, 338. The tests A and B under Rhythm follow Titchener quite closely.

blows of a hammer, etc. In some such way these persons will, involuntarily, manifest their subjective transfer of the series of uniform sounds into a succession of similar accented groups. If they are at all keen, introspectively, they will at the same time report that every second, third or fourth sound, as the case may be, is stressed, and forms the initial member of a group of two, three or four sounds, and that the interval between groups is longer than that between the individual sounds. Other persons will hear only the succession of similar sounds, each like the preceding, and each separated from the others by a uniform time interval; and until some more or less direct suggestion of rhythmisation is made to them, they will not transpose the uniform series into rhythmical groups. The former class possess good rhythmical sense; the latter are more or less deficient. Both classes are to be examined further by the tests with suggestion, described in the next section.

B. *Rhythmisation with Suggestion.* The first test under this head is to determine the individual's normal rhythmical grouping, *i. e.*, the grouping into which sounds most easily fall when he comes to group them rhythmically. If the preceding tests show many instances of subjective rhythmisation, the person's normal rate will already be evident. Even in such cases, however, it will be well to go through these additional tests in which some indirect suggestion is made to bring out subjective grouping more prominently. Suggestion may be effected as follows: The subject may be asked to count the beats of the metronome, and asked later how he counts them; or if he said the sounds were "like a clock," he may be asked if they differed in quality and intensity as the ticks of a clock do, and in this way his attention may be directed toward grouping the sounds (Bolton). In case these suggestions fail to evoke groupings, the subject should be asked outright to note if the beats do not group themselves, as he listens to them. By this means, the person may be led to make subjective groupings. The results will show what grouping (usually the two-group) is favored.

The second division of this test measures the person's ability to throw a uniform series of sounds at will into any grouping desired. The series already given should be repeated, and others given at slower rates, while, in each series, the person suggests to himself arrangements of the sounds into groups of 2, 3, 4, 5, 6, 7, or 8 sounds each. The 2-group, and multiples of it up to 4, 6, and sometimes 8, will usually be found easiest to suggest; the 3-group, fairly easy; the 5-group, usually difficult; and the 7-group, extremely difficult. Imagery will be found useful in assisting suggestion: a pendulum for 2-groups; a triangle, perhaps, for 3-groups; a square for 4-groups, etc.

The first part of this experiment examines the individual's capacity for involuntary rhythmical grouping, and this, with the test on rhythmisation under suggestion, will throw light on the rhythmical sense of the person tested. Those who, in the first test, give no evidence of subjective rhythmisation and in the latter test are not able by any sort of suggestion to effect a grouping of the sounds, may be considered deficient in a sense of rhythm.

With the above tests several supplementary tests of the sense of rhythm may be mentioned. In each, the series of uniform sounds is replaced by a series in which piano notes are used instead of metronome clicks, and in which a periodic variation in duration, time interval, or quality is introduced.[1]

(a) Varied duration. The following series of notes is played with equal intensity upon the same piano-key: quarter—half—quarter, quarter—half—quarter, etc.; quarter—quarter—quarter—half, quarter—quarter—quarter—half, etc.; quarter—half—quarter, quarter, etc. The longer sound ordinarily appears more intensive than the rest, and the interval following is apparently longer than the other intervals, thus giving rhythmical grouping, with the lengthened note forming the first member of the groups.

(b) Varied time interval. Series with periodically recurring lengthened intervals are played on one digital: quarter—quarter—quarter—rest, quarter—quarter—quarter—rest, etc. The sound followed by the rest appears louder than the others and begins rhythmical groups.

(c) Varied quality. Two, three or more notes *e. g.*, $c^1$—$e^1$ or $c^1$—$d^1$—$d^1$—$e^1$, are repeated over and over in a series, with equal intensity and at equal intervals, thus: $c^1$—$e^1$—$c^1$—$e^1$—$c^1$—$e^1$, etc., and $c^1$—$d^1$—$e^1$—$c^1$—$d^1$—$e^1$—$c^1$—$d^1$—$e^1$, etc. Similarly for a group of four notes. Such a series will usually be arranged in 2-groups, 3-groups, or 4-groups, each composed of the repeated notes.

As already stated, musical compositions are ordinarily written on the basis either of the 2-beat or of the 3-beat rhythmical grouping. A test is suggested of playing a duple time composition, and another in triple time. The terms "waltz," "two-step," "march," etc., carry a rhythmical suggestion, and should not be used in connection with the test. Simply submit two musical "compositions," and ask for a judgment as to the rhythmical grouping of each. This last experiment can be performed with the piano or some other musical instrument, or with a phonograph, gramophone, or even a good music box.

[1] Titchener: *op. cit.*, Vol. I, pt. 2, 348-9.

C. *The Perception of Musical Phrases and Sentences.* If a person fully appreciates music, a composition becomes for him an organized movement in which larger parts, musical sentences, follow each other in rhythmical recurrence, and, at the same time, present within themselves smaller recurring parts, musical phrases. The "phrase" usually includes 4-8 measures, and the "sentence" 8-16 measures, and a smaller grouping, of two measures, is recognized. The perception of this form or structure of music is perhaps the largest factor in æsthetical musical perception, and yet it can be tested only inadequately.

The test for phrasing will consist in presenting one or more simple and rather distinctly phrased compositions, and requiring a judgment as to the number of large divisions of which it is composed. It will be necessary to use music with which the persons are unfamiliar. Old and unusual accompaniments might be selected. The music of a foreign national anthem, if unfamiliar, might be used. It may be found necessary to preface the test with a statement of the nature of musical phrases and sentences, and quite possibly with illustrations from actual music. The piano, organ, or phonograph can be used in the tests.

IV. *Affective Reaction to Music.* The keen enjoyment which some persons secure from music, and the comparative indifference with which it is regarded by others, are evidence of individual differences in the affective reaction to music. Here, then, is a factor in musical capacity which can be examined apart from the factors already considered.

A test of capacity for musical enjoyment can best be made by using simple tonal combinations rather than complex musical compositions, as the degree of pleasure derived from the latter is so largely influenced by training, familiarity with the particular composition and with music in general, that it would be difficult to interpret the results. The "Star Spangled Banner," for example, would appeal to the average American in a degree beyond its musical merits, while the "Marseillaise" would be devoid of associations. Accordingly, simple tests with intervals and chords should first be used, though they may properly be supplemented by tests with whole musical compositions. The simpler tests are further advisable, since we have clear evidence, based on our musical inheritance, as to the normal relative pleasantness and unpleasantness of the different intervals and chords; thus, in such tests, a standard is already provided by which the individual's result can be evaluated.

The following tests are suggested:

*A.* Harmonic intervals, *i. e.*, combinations of two simultaneous tones, should be submitted for paired comparison, and

judgment asked as to which is the more pleasant or unpleasant.[1] Judgment should be passed on the following intervals: octave and 5th; octave and major 3rd; major 2nd and major 3rd; 7th and major 3rd; 5th and major 3rd; octave and major 6th; 5th and major 6th; major 3rd and major 6th; octave and minor 3rd; octave and minor 6th; 7th and major 6th; 7th and major 3rd. The results should show a preference for the 3rds and 6ths.

After these tests on harmonic intervals, the same combinations may be repeated as melodic intervals, and judgment again passed as to pleasantness and unpleasantness. The first test in the series just suggested would then be given as follows: The tones in the octave $c^1$—$c^2$ should be given in close succession, followed after 2 or 3 seconds by the two tones of the 5th ($c^1$ and $g^1$); and judgment asked as to which combination is more pleasant. Similarly with the octave and major 3rd, etc. In this series, also, judgments will probably favor the 3rds and 6ths.[2] If but one of the tests is to be given, the first, *i. e.*, the harmonic, is to be preferred.

*B.* The larger simultaneous combinations of tones—chords—should also be compared as to their relative pleasantness. The method, as before, is that of paired comparison.

Material for tests can be secured by arranging consonant and dissonant triad chords in pairs for judgment.[3] The following list is suggested for testing, and after each pair of chords, the chord is indicated which should normally be considered more pleasant. The test can be performed on a piano or organ.

| | | |
|---|---|---|
| $c^1 - e^1 - g^1$; | $c^1 - e^1 - f^1$ | 1st preferred. |
| $c^1 - d^1 - g^1$; | $c^1 - e^1 - g^1$ | 2nd " |
| $c^1 - e^1\flat - g^1$; | $c^1 - a^1 - g^1$ | 1st " |
| $c^1 - e^1\flat - f^1$; | $c^1 - f^1 - a^1$ | 2nd " |
| $c^1 - e^1\flat - b^1\flat$; | $c^1 - f^1 - a^1$ | 2nd " |
| $c^1 - e^1 - a^1$; | $c^1 - e^1\flat - a^1$ | 1st " |
| $c^1 - e^1 - a^1\flat$; | $c^1 - e^1\flat - a^1\flat$ | 2nd " |
| $c^1 - g^1 - a^1$; | $c^1 - f^1 - a^1$ | 2nd " |
| $c^1 - e^1 - g^1$; | $c^1 - b^1\flat - a^1$ | 1st " |
| $c^1 - a^1\flat - b^1\flat$; | $c^1 - e^1\flat - g^1$ | 2nd " |
| $c^1 - e^1 - g^1$; | $c^1 - e^1 - g^1\sharp$ | 1st " |

Each pair contains one consonant chord and one dissonant chord, and the term preferred is in each case given as the consonant combination. Consonance in chords in general implies pleasantness. Persons tested should be given no information regarding the chords; but should be asked simply to report each time the chord which is the more pleasant.

---

[1] Stumpf, as already noted, used this as one test for unmusical ears, and found that, with unmusical persons, atrocious combinations of tones, as long as directly neighboring tones were not included, were usually not considered unpleasant. *Tonpsych.*, II, 158.

[2] Helmholtz: *Sensations of Tone*, 1895, 370: "The most attractive of the intervals, melodically and harmonically, are clearly the 3rds and 6ths."

[3] Helmholtz, *ibid.*, 212, gives a table of consonant and dissonant chords.

*C.* A final test for affective appreciation of music may be made by presenting in succession different musical compositions, and asking for preferences. So far as the persons are conscious of reasons for preference, these should also be stated. This test should at least show the presence or absence of bare liking for one piece of music more than for another. Comparative judgment may be passed on such pieces as :

1. The Marseillaise and Die Wacht am Rhein.
2. Swannee River and some modern popular negro music.
3. A waltz and a two-step, which are, as nearly as possible, equally familiar, and of equal merit as musical compositions.
4. Phrases selected from the work of different composers, or from different music by the same composer.
5. Two whole masterpieces of musical composition. If possible these should be actually performed in the presence of the persons tested, care being taken to avoid suggestions as to the identity or merits of the composition, or other remarks which would influence the judgment given. In the absence of suitable instruments, the phonograph will be a second-rate substitute.

In all these affective tests, it is important that the subject record not only whether the effect is pleasant or unpleasant, but, roughly at least, the strength of the reaction, as very pleasant, slightly pleasant, indifferent, slightly unpleasant, very unpleasant.

By these three tests, an examination of an individual's affective reaction to music is possible, first as to whether such a reaction is present, and secondly as to its strength. The first two tests will show whether the elements of musical composition, the intervals and chords, are heard with varying degrees of pleasantness and unpleasantness, and whether the preferences for certain combinations agree with the standards of affective value ordinarily received. The third test will disclose something of the nature and strength of the affective reaction to whole musical compositions.

V. *Miscellaneous Tests and Methods.* A small percentage of persons possess an absolute memory for pitch.[1] They are able on hearing a tone, to assign it its proper place in the scale, as $g^1$, $c^1$, g. To test absolute-pitch memory, the persons should be placed so that they cannot see the key-board of the piano, and then required to name single piano notes which are sounded in irregular order.

[1] F. C. French (*Psych. Rev.* IX, 1902, 40, *Mental Imagery of Students*) reports that 14 out of 117 students (12%) declared that they had an absolute memory for pitch. This percentage is unusually high, and it seems likely that, with actual experimentation, it would have been considerably decreased.

Vocal repetition of a tone given by the piano, or previously sung, tests at once the perception for tones and the ability to reproduce them vocally, and is not to be regarded as a test simply of accurate auditory perception of tone. There is a possible source of error, too, in the evaluation of the reproduced tone by the examiner. Nevertheless, this test will be found useful. Stumpf, it will be remembered, includes it in his list.

Another method, often used as an exercise in musical instruction, and capable of adaptation as a test, is to require the observer to sing notes as they are pointed out on a scale arranged conveniently on a chart or blackboard. The notes may be called for in regular succession, up and down the scale, or with larger intervals, *e. g.*, *do*, *me*, *sol*, etc. This is primarily a test of correct vocalization, and audition is concerned only so far as accurate hearing is always implied in accurate singing.

Monroe used three musical tests on small children, 2 to 6 years old, which may find application elsewhere under certain circumstances: (1) Ability to learn the scale; (2) Ability to remember the scale after two weeks; (3) Ability to learn kindergarten rote-songs and reproduce them after two weeks.[1]

Considerable information regarding musical capacity can be obtained from a series of questions put to the persons tested. Parents of children should also be questioned. The following questions are suggested:[2]

1. Musical education and training: training in public schools? private lessons, instrumental and vocal, when begun and how long continued? musical environment in childhood? attendance at operas, concerts and recitals? interest in theory of music? interest in musical composers and history of music? do parents or other members of the family show special musical ability or interest?

2. Enjoyment of music: in performance and in listening? music lessons in school? favorite instrument, voice, composer and composition? pleasure in marching to band music? pleasure in dancing? is part of enjoyment based on knowledge of music?

---

[1] W. S. Monroe: *Ped. Sem.*, X, 1903, 144. His results are given in percentages. Of the boys, 34%, and of the girls, 58%, learned the scale. After a fortnight, with three intervening reviews, 29% of the boys and 40% of the girls could reproduce it. Kindergarten songs were remembered after a fortnight by 50% of the boys and 63% of the girls. On the basis of their individual reaction to the tests, and information furnished by parents, 27% of the boys and 59% of the girls are credited with special taste for music.

[2] The author is indebted for suggestions in certain questions to S. E. Sharp, *Individual Psychology, Amer. Jour. of Psych.*, X, 1899, 359 and 369.

3. Skill in music: vocal and instrumental? taking part in chorus singing, as in church? solo singing? playing in public? ability to carry an air? to reproduce it in one's head? by whistling? by singing?

In summarizing, it may well be reiterated that appreciation of music as an æsthetic experience is more than the factors which have been mentioned. Nevertheless, on the results of the tests and of the questions suggested, an approximate measure of the individual's "musical capacity" may be safely based. The capacity of the person to discriminate differences in pitch and in chords, to appreciate rhythm and to experience an affective reaction from music, can all be ascertained by such methods as we have described. The tests outlined will serve to disclose persons of unusual musical gifts, and to call attention to the desirability of training them; and, as well, to indicate those who are tone-deaf or otherwise incapacitated by nature from responding to musical training. A final word of warning might be said, that no child be considered hopelessly unmusical until given an opportunity for musical training. In musical capacity, much depends on practice.

## CHAPTER V.

### Diagnostic Tests of Hearing.

This section considers those tests of audition which may be employed in the diagnosis of auditory disease and defect. We pass here from the domain of psychology and mental anthropometry to that of pathology; or, more precisely, to a particular field of pathology, otology. Some of the tests already treated have otological value; others belong more exclusively to this field. Such tests are of interest from a psychological standpoint, and should be treated, even if but briefly, to make our systematic presentation complete. The tests we shall consider are the following:

(1) Diagnostic Speech Tests.
(2) Integrity of the Tonal Scale.
(3) Determination of Deafness.
(4) Pitch Difference of the Two Ears.
(5) Special Diagnostic Tests:
    (a) Weber, (b) Schwabach, (c) Gellé, (d) Rinne.

In general, it may be said that the following empirical facts lie at the basis of methods of diagnostic examination: The nature of auditory perception with diseased and defective organs varies (1) with different forms of stimulation, as with the tuning fork and the human voice; (2) at different parts of the tonal scale, as for high and low notes; and (3) with differ-

ent kinds of conduction (aerial or osseous). One or another of these facts appears in the tests to which we now pass.

I. *Diagnostic Speech Tests.* The use of speech tests in otological diagnosis depends upon the establishment of correlations between deafness or diminished activity for particular elements of speech on the one hand, and particular derangements of the auditory organ, on the other. American and English otologists have apparently done little in this field, and we can simply indicate the work done by German otologists, and call attention to the need for further study of the problem. Oscar Wolf seems to have made the first systematic application of speech to aural diagnosis. He tested auditory organs, which were variously diseased, as to the comparative ease of perception for the three classes of consonants, low, medium, and high, referred to in our chapter on speech tests. His material included these speech elements pronounced separately, and combined in words. As the result of his investigations, he correlated defective hearing for certain speech-sounds with various pathological conditions of the ear. Some of these correlations are as follows: poor perception of s-sounds is found with obstruction in the sound-conducting apparatus; loss of deep, lingual r with defects of the ear drum; and loss of the f-sound, with labyrinthine disease.[1] An important general principal stated by Wolf is that deep tones are heard with difficulty when there is disturbance of the sound-conducting apparatus. Bezold confirms this and states that, with number-word tests, loss of perception for 100 (hundert) is particularly characteristic of disturbance in the sound-conducting apparatus. Bezold gives other correlations, among them these: Poor perception of 5 (fünf) with inflammation of the middle-ear, 7 (sieben) and words with high consonants, with labyrinthine diseases. Bezold does not test perception for all speech sounds, but only for those low and high sounds which he has found most often defectively perceived.[2] Such diagnostic evidence is never entirely conclusive; it simply supplements other methods of diagnosis, and as a preliminary test gives valuable hints for further procedure. German otologists are by no means entirely agreed as to the significance of loss of perception for the various speech elements; but the facts already mentioned indicate the possibility of establishing a tentative differential diagnosis upon the results of speech-tests. American and English otologists might soon build up a similar system for English words by testing patients with definite words like the number series, and correlating defective hearing with the pathological conditions present.

[1] Wolf, quoted by Politzer, *Ohrenheilkunde*, 116.
[2] Bezold: *Funktionelle Prüfung*, 207 ff.

II. *Integrity of the Tonal Scale.* Some aural diseases manifest themselves by loss of certain tonal sensations: *e. g.*, low tones, as Wolf observed, are poorly heard if the sound-conducting apparatus is disturbed, and in certain other diseases the perception of high tones may be affected. Moreover, cases are not infrequent in which there is loss of hearing for parts of the scale, *i. e.*, tonal gaps are present; or, retention of hearing for isolated parts, tonal islands, above and below which there is deafness for tones. Tests to determine the condition of tonal perception at various points along the scale are therefore necessary. Politzer recommends for such a test the use of at least three forks: c (128 vib.), $c^2$ (512 vib.), and $c^4$ (2048 vib.).[1] Bezold has devised apparatus consisting of eleven forks with adjustable weights, two high adjustable organ pipes, and the Galton whistle.[2] This comprises his "continuous tonal series" with which it is possible to give tonal stimuli at every semitone interval throughout the scale, from a lower limit of eleven vibrations in the second to the upper limit of audibility for tones.

In interpreting the results of such tests, it is commonly stated that, with air conduction, impaired perception of high tones indicates disease of the internal ear, and of low tones, disturbance of the middle ear; but this is only conditionally valid. The loss of perception for high notes can be taken as only partially conclusive, and must be confirmed by other diagnostic evidence.[3] The loss of perception for low tones, however, is quite trustworthy evidence of middle-ear disturbance.[4] In any event, the results of both tests, and especially the test with low tones, are extremely valuable in diagnosis.

Bezold has made an application of the continuous tonal-series apparatus in tests of deaf mutes, who frequently retain audition for certain pitches and lose it for others, so that tonal gaps and islands are present. The accurate knowledge of what pitches are still audible may, as he suggests, be turned to account in the education of these defectives. The results promise, as well, to throw light on the diagnosis of functional

[1] Politzer: *Ohrenheilkunde*, 111.

[2] The apparatus is made by Edelmann, of Munich, and described by Bezold, *op. cit.*, 123 ff., 217 ff., and 229 ff.

[3] Politzer: *op. cit.*, 112.

[4] Politzer notes some exceptions to it (*op. cit.*, 112), but Bezold accepts it as final (*op. cit.*, pp. 219, 222); and Lucae says that if low tones are heard normally, no essential disturbance to the sound-conducting apparatus can be present (*Arch. f. Ohrenheilk.*, XV, 280). Bezold has accordingly concluded that the peculiar physiological function of the chain of ossicles is the transmission of vibrations of low rates, corresponding to the deeper tones (*op. cit.*, 221).

auditory disturbances, and they have an important bearing on the acceptance or rejection of various theories of hearing.[1]

In addition to these qualitative tests, which examine the condition of perception at various points along the whole tonal scale, many otologists use a temporal test of similar comprehensiveness which measures the ringing-off time for forks with air conduction. This temporal test is applied at intervals of one octave or less up and down the scale. Such tests require considerable time and are only approximately accurate, since it is impossible to secure a constant intensity in successive tests, and difficult for the patient to tell just when a fork is no longer audible; yet they furnish indispensable assistance in diagnosis, as Bezold says; particularly in the manifold forms of cochlear disturbance.[2]

III. *Determination of Deafness.* Deafness may be of two forms: (1) a deafness for speech, or relative deafness, in which there may persist a very considerable ability to hear tones, and (2) absolute deafness, in which neither speech nor tones can be heard.[3]

If the deafness, whether relative or absolute, exists in both ears, its determination presents little difficulty. Each ear is examined in turn, the other, meantime, being stopped. Speech tests are used to determine relative deafness; and individual forks, or better those of the tonal series, to determine absolute deafness. The former test is practically the acuity test with speech; the latter aims to discover what tonal qualities, if any, are still sensed.

One-sided deafness presents more difficulties. If the case is relative deafness, results must be checked by Dennert's method, which, it will be recalled, consists of closing both ears and repeating the test words heard, to be sure that they were originally perceived by the ear under examination and not by the normal ear. If the case is absolute one-sided deafness, the tonal series apparatus is used: for the lower tones up to $c^1$, Bezold says that it is not necessary even to close the normal ear to prevents its functioning; in the next octave, $c^1$ to $c^2$, it is only necessary to close the normal ear with the finger to secure complete isolation. Above $c^2$, it is impossible to secure isolation even with closure, as the high notes are very penetrating. Perception for these high notes, however, can be tested indirectly. Bezold determined the ringing-off times for the various

[1] Bezold (*op. cit.*, 229-240, also printed in *Zeits. f. Psych. u. Phys.*, XIII, 1896, 161) gives results of examinations of deaf-mutes and their bearing on the Helmholtz theory of audition. *Cf.* also *Funkt. Prüf.*, 224.

[2] Bezold: *op. cit.*, 223.

[3] Bezold: *op. cit.*, 224 ff.

high forks with persons from whom the cochlea had been removed on one side, and who were accordingly entirely deaf on that side. The fork was sounded on the deaf side, and ringing-off times taken. Thus conduction-norms were secured, to which the results of high-fork tests with one-sided absolute deafness can be referred.

IV. *Pitch Difference Between the Two Ears.* If the same tonal stimulus is presented successively to the right and the left ear, it is normally perceived as of slightly higher pitch by the one ear than by the other. The difference is ordinarily small, from a few vibrations to one-fourth of a tone, and the right ear is usually the higher. This difference is a normal phenomenon, and the determination of its magnitude is a classical problem in experimental psychology.[1] The method may be briefly stated: Two forks, of the same objective pitch, give slightly different pitches, when heard, one at the right ear, the other at the left ear. With a bit of wax, the subjectively higher fork is flatted to subjective equality when heard as before. The difference in vibration rates is then measured in beats over a resonance bottle, and is transferred into an expression of the pitch difference of the two ears.

Aside from this normal pitch difference between the two ears, there are pathological cases (*diplacusis binauralis*) in which a difference of one-half or a whole tone, or even a musical third or more, appears. The method given for determining the small normal pitch difference would have to be superseded in such cases by the use of a series of forks giving differences of from one-fourth of a tone up. Another pathological phenomenon involving pitch differences in audition is double hearing (*paracusis duplicata*); in this disturbance, two distinct auditory perceptions arise from a single stimulus. It may be confined to speech or to tones, or may extend to both; but it is most common with tones. These disturbances usually arise from catarrh and inflammation of the middle ear; though many cases are explained by reference to the basilar membrane. Both demand the attention of the aurist.[2]

V. *Special Diagnostic Tests.* Otological diagnosis employs a series of special tests, based, in general, upon the alterations due to disease in the normal relation between bone-conduction and air-conduction of sound waves. For the normal ear, a fork sounds longer before the meatus with air-conduction than it does with osseous conduction—as when set upon the mastoid process of the skull. With certain diseases of the ear, this re-

[1] Titchener: *op. cit.*, Vol. I, Pt. I, 38.

[2] For both these phenomena, see Politzer, *op. cit.*, 557-558; Gruber: *Diseases of the Ear*, 137-140; Stumpf: *Tonpsych.*, I, 266 ff., 273-8.

lation is altered.[1] We can do but little more than state the tests that are based upon this principle, and explain their general significance. The actual interpretation of the results given by the tests involves so many interdependent factors and the recognition of so many exceptions, that it would be beside our purpose to make a systematic statement of the rules of auditory diagnosis. That belongs to the science of otology. The tests we shall describe are Weber's, Rinne's, Gellé's and Schwabach's.

A. *Weber's Test.* E. H. Weber discovered that if a vibrating tuning fork is placed upon the median line of the skull, while one ear is closed, as by the finger, the sound is localized in the closed ear.[2] If there is an obstruction in the middle ear, the sound of the fork is similarly localized in the obstructed ear. This localization of sound in the ear which is stopped is explained as due (1) to increased resonance of the outer aural passages (Weber); (2) to transfer of vibrations from the skull bone to the air of the meatus where they are reflected by the obstruction upon the tympanic membrane; and (3) to the altered tension of this membrane and the ossicles (Politzer). The lateralization with middle ear diseases is explained by Bezold as due to increased rigidity of the ossicles and their ligaments, which favors the transfer of vibrations from the bones of the skull to the ossicles. This rigidity interferes, at the same time, with the delicate equipoise of the sound-conducting apparatus and hinders the usual passage of sound vibrations by way of air-conduction and the tympanic membrane. The method ordinarily employed consists, as stated, in placing a fork on the median line of the skull and observing whether the sound is localized in one ear. The general interpretation of results is as follows: With but one ear affected, the localization of sound in it points almost unmistakably to an obstruction in the sound-conducting apparatus; in disturbance affecting both ears, but unequally, localization in the more seriously affected ear points to the same conclusion. On the contrary, localization of sound in the normal or less affected ear is of slight significance unless supported by other diagnostic evidences. Bezold recommends the use of the unweighted $a^1$-fork for this and the other tests with bone-conduction.

B. *Rinne's Test.* This test compares directly the difference in ringing-off time for air and bone conduction.[3] The fork is set with the base on the mastoid process and allowed to ring-

[1] See Bezold, *op. cit.*. 44 ff., for a statement of theories of bone and air-conduction.

[2] *De Pulsu, Auditu et Tactu*, 1834.

[3] *Beiträge zur Phys. d. menschl. Ohres, Prager, Vierteljahrsschrift*, i, 1855, 71; ii, 45, 155.

off; it is then transferred immediately to a position before the ear, where, if the ear be normal, it again becomes audible and continues so for some time. In such a case, the test is said to fall out positively (Rinne +, the normal result); and the result is expressed by the temporal excess of air-conduction over bone. If the contrary condition is true, and bone-conduction is better than air, the test is performed in the reverse order, first with air and then with bone-conduction, and the temporal excess is measured as before. In this case, the test falls out "negatively" (Rinne—, defective). The normal result with the $a^1$-fork is positive 30 seconds, *i. e.*, air-conduction lasts 30 seconds longer than bone (under Bezold's conditions, of course). The normal excess becomes less in old age.[1] With both ears affected by disease, and the acuity of the two ears not widely different, negative Rinne indicates a disturbance of the sound-conducting apparatus. With one ear diseased, the negative result may occur, though the conducting apparatus be intact. A normal, or nearly normal, positive Rinne, when hearing acuity is considerably reduced, taken with other diagnostic indications, excludes any affection of the conducting apparatus.[2]

C. *Gellé's Test.*[3] A vibrating fork is placed on the skull while the pressure of air against the tympanic membrane is increased by means of an air-ball fastened to a tube, or Siegle's speculum. In a normal ear, the tone is decreased in intensity coincident with increased pressure against the tympanic membrane; this is explained as due (1) to increased pressure in the cochlea by reason of forcing the ossicles inward, and (2) to changed conditions in the sound-conducting apparatus. If there is an obstruction in the conduction-apparatus, there is no decrease in the intensity of the tone heard; on the contrary, if the conduction-apparatus is intact, and the cochlea affected, there is the decrease in intensity noted in normal ears. This is Gellé's general rule; but in actual use, as with all these tests, many complicating factors must be taken into account. The test has practical difficulties, too, in the quick dying-off of the tone, and the unreliable statements of patients regarding changes in its intensity. It is chiefly employed in cases of severe disturbances of hearing, where other methods do not give definite results. (Politzer, p. 126.)

D. *Schwabach's Test.* This test compares directly bone-conduction for the diseased ear with that of a normal ear. A

[1] Bone-conduction does not decrease absolutely in old age, but runs parallel with the decrease in acuity of audition. Bezold: *op. cit.*, 166.

[2] Reference must be had to Bezold, *op. cit.*, 55-122, especially p. 116, and to Politzer, *op. cit.*, 124-5, for complete statements of the interpretation to be given results.

[3] *Precis des maladies de l'oreille*, Paris, 1885, 339-40.

sounding fork is placed on the mastoid process of the ear to be examined, and when it is no longer heard, it is transferred to the mastoid process of the examiner's normal ear, and the excess time noted; or, if necessary, the procedure is reversed, and the excess for the diseased ear measured. In the first case, the diseased ear has sub-normal bone-conduction; in the latter, super-normal. The general principle for interpreting results is that disease of the external or middle ear is accompanied with super-normal (prolonged) bone-conduction; and disease of the internal ear, with sub-normal (shortened) bone-conduction. The test is not valid by itself. Super-normal Schwabach with negative Rinne, however, indicates an obstruction in the sound-conducting apparatus; sub-normal Schwabach with positive Rinne and high-grade deafness for speech indicates a cochlear disease. Schwabach's test is not reliable with one-sided affection of hearing, or with double-sided disease in which there is much difference in the degree of disturbance on the two sides.[1]

## General Summary.

The chief problems of the paper[2] may be summarized as follows:

I. The definition of mental tests in general and their relation to experimental psychology. The position was taken that there is a very clear distinction between tests and psychological experiments proper, in that tests are rapid, approximate measurements applied to miscellaneous persons for practical purposes; while psychological measurements are made with refined methods, on practised observers, for scientific ends. A psychology of individual variations will doubtless be developed, but it will depend, the writer believes, on investigations of the latter rather than of the former sort.

II. The literature of mental tests. The movement was found to have arisen in Galton's anthropometric measurements. A psychological trend was imparted to it by Cattell, Jastrow, and the Committee of the American Psychological Association, so that mental tests and statistical manipulation of the results of measurements came to be regarded as a new and promising method of psychological investigation. This, the writer holds is a wrong conception. Titchener's article of 1893 called attention to the real separation of psychology and mental tests. The distinction which he made has come gradually to be accepted. This paper urges explicit recognition of the distinction, and acknowledgment that mental tests are not

[1] Politzer: *op. cit.*, 121.

[2] The first part of the paper, Chs. I-III, appeared in this *Journal* for Jan., 1904.

a part of psychology, but, as has long ago been said, are properly mental anthropometry. Such a distinction ascribes to tests only practical extra-psychological ends, and this is the standpoint taken by the present writer. The problem in hand becomes then, the description of anthropometric tests of audition. The problem naturally falls into two parts, general hearing tests and tests of musical capacity. To these, diagnostic tests were added to make a complete treatment of the subject.

III. General hearing tests were treated in two parts: (A) Speech tests, in which, after an examination of the problem, whisper-tests, employing series of number-words, were selected as the most feasible, and an improved method of testing was suggested; and (B) tests employing mechanical sounds, the watch, acoumeter, audiometer, and forks. With the latter division were considered tests to determine the lower and upper limits of tonal audibility. In general, acuity of hearing should be measured by speech tests, though mechanical tests form a possible substitute when convenience requires their use. In the appendix to this section, there are presented the results of experiments in the use of speech in auditory testing. The traditional method of extreme range in speech tests is shown to be unreliable, and a new method of degree of accuracy is proposed.

IV. Tests of musical capacity. Musical perception is too complicated to be tested directly as a unitary experience. Several sub-factors, however, were found to admit the application of tests: I. Pitch discrimination, measured by the adjustable pitch pipe, forks, or roughly, by the piano; II. Simultaneous tonal combinations; III. Rhythm and phrasing; IV. The affective reaction upon music; V. Miscellaneous methods, including a series of questions. The results of tests upon all these factors, taken together, give a basis for an approximate judgment of musical capacity.

V. Diagnostic tests. This section treated tests used in otological practice to localize and identify aural disturbances and diseases. The tests considered include: I. Diagnostic speech tests. II. Tests for integrity of tonal scale; III. Deafness; IV. Pitch difference between the two ears; V. Special tests devised by Weber, Rinne, Gellé, and Schwabach.

It is unnecessary to say that most of the diagnostic tests are of significance to the practising aurist alone. The tests of general practical value which have been stated are those of Chapters III and IV, anthropometric in nature, measuring the two important auditory functions, perception of speech and perception of music. These two sets of tests are submitted as contributions to what the writer believes should be termed mental anthropometry.